K. Carlson

PAGE: 1

RAW SEQUENCE LISTING
PATENT APPLICATION   US/09/207,161

DATE: 03/22/2000
TIME: 11:50:12

INPUT SET: S35114.raw

#10
4/13/2000

> This Raw Listing contains the General
> Information Section and up to the first 5 pages.

```
1                                    SEQUENCE LISTING
2
3    (1)    General Information:
4                                                  ENTERED
5         (i) APPLICANT: Hillman, Jennifer L.
6                        Goli, Surya K.
7
8         (ii) TITLE OF INVENTION: NOVEL HUMAN INTEGRAL MEMBRANE
9                   PROTEIN
10
11        (iii) NUMBER OF SEQUENCES: 4
12
13        (iv) CORRESPONDENCE ADDRESS:
14           (A) ADDRESSEE: Incyte Pharmaceuticals, Inc.
15           (B) STREET: 3174 Porter Drive
16           (C) CITY: Palo Alto
17           (D) STATE: CA
18           (E) COUNTRY: USA
19           (F) ZIP: 94304
20
21        (v) COMPUTER READABLE FORM:
22           (A) MEDIUM TYPE: Diskette
23           (B) COMPUTER: IBM Compatible
24           (C) OPERATING SYSTEM: DOS
25           (D) SOFTWARE: FastSEQ for Windows Version 2.0
26
27        (vi) CURRENT APPLICATION DATA:
28           (A) APPLICATION NUMBER: 09/207,161
29           (B) FILING DATE:
30           (C) CLASSIFICATION:
31
32        (vii) PRIOR APPLICATION DATA:
33           (A) APPLICATION NUMBER: US/08/791,338
34           (B) FILING DATE:
35
36        (viii) ATTORNEY/AGENT INFORMATION:
37           (A) NAME: Billings, Lucy J.
38           (B) REGISTRATION NUMBER: 36,749
39           (C) REFERENCE/DOCKET NUMBER: PF-0208 US
40
41        (ix) TELECOMMUNICATION INFORMATION:
42           (A) TELEPHONE: 415-855-0555
43           (B) TELEFAX: 415-845-4166
44
45
46          (2) INFORMATION FOR SEQ ID NO:1:
```

INPUT SET: S35114.raw

```
47
48          (i) SEQUENCE CHARACTERISTICS:
49              (A) LENGTH: 266 amino acids
50              (B) TYPE: amino acid
51              (C) STRANDEDNESS: single
52              (D) TOPOLOGY: linear
53
54          (vii) IMMEDIATE SOURCE:
55              (A) LIBRARY: NEUTGMT01
56              (B) CLONE: 632664
57
58          (xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:
59
60   Met Val Lys Val Thr Phe Asn Ser Ala Leu Ala Gln Lys Glu Ala Lys
61    1               5                  10                  15
62   Lys Asp Glu Pro Lys Ser Gly Glu Glu Ala Leu Ile Ile Pro Pro Asp
63               20                  25                  30
64   Ala Val Ala Val Asp Cys Lys Asp Pro Asp Asp Val Val Pro Val Gly
65               35                  40                  45
66   Gln Arg Arg Ala Trp Cys Trp Cys Met Cys Phe Gly Leu Ala Phe Met
67        50                  55                  60
68   Leu Ala Gly Val Ile Leu Gly Gly Ala Tyr Leu Tyr Lys Tyr Phe Ala
69    65                  70                  75                  80
70   Leu Gln Pro Asp Asp Val Tyr Tyr Cys Gly Ile Lys Tyr Ile Lys Asp
71               85                  90                  95
72   Asp Val Ile Leu Asn Glu Pro Ser Ala Asp Ala Pro Ala Ala Leu Tyr
73               100                 105                 110
74   Gln Thr Ile Glu Glu Asn Ile Lys Ile Phe Glu Glu Glu Glu Val Glu
75               115                 120                 125
76   Phe Ile Ser Val Pro Val Pro Glu Phe Ala Asp Ser Asp Pro Ala Asn
77        130                 135                 140
78   Ile Val His Asp Phe Asn Lys Lys Leu Thr Ala Tyr Leu Asp Leu Asn
79    145                 150                 155                 160
80   Leu Asp Lys Cys Tyr Val Ile Pro Leu Asn Thr Ser Ile Val Met Pro
81               165                 170                 175
82   Pro Arg Asn Leu Leu Glu Leu Leu Ile Asn Ile Lys Ala Gly Thr Tyr
83               180                 185                 190
84   Leu Pro Gln Ser Tyr Leu Ile His Glu His Met Val Ile Thr Asp Arg
85               195                 200                 205
86   Ile Glu Asn Ile Asp His Leu Gly Phe Phe Ile Tyr Arg Leu Cys His
87        210                 215                 220
88   Asp Lys Glu Thr Tyr Lys Leu Gln Arg Arg Glu Thr Ile Lys Gly Ile
89    225                 230                 235                 240
90   Gln Lys Arg Glu Ala Ser Asn Cys Phe Ala Ile Arg His Phe Glu Asn
91               245                 250                 255
92   Lys Phe Ala Val Glu Thr Leu Ile Cys Ser
93               260                 265
94
95          (2) INFORMATION FOR SEQ ID NO:2:
96
97          (i) SEQUENCE CHARACTERISTICS:
98              (A) LENGTH: 1114 base pairs
99              (B) TYPE: nucleic acid
```

INPUT SET: S35114.raw

```
100            (C) STRANDEDNESS: single
101            (D) TOPOLOGY: linear
102
103        (vii) IMMEDIATE SOURCE:
104            (A) LIBRARY: NEUTGMT01
105            (B) CLONE: 632664
106
107        (xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:
108
109    GCCGCCTCTG CCGCCGCGGA CTTCCCGAAC CTCTTCAGCC GCCCGGAGCC GCTCCCGGAG      60
110    CCCGGCCGTA GAGGCTGCAA TCGCAGCCGG TGAGCCCGCA GCCCGCGCCC CGAGCCCGCC     120
111    GCCGCCCTTC GAGGGCGCCC CAGGCCGCGC CATGGTGAAG GTGACGTTCA ACTCCGCTCT     180
112    GGCCCAGAAG GAGGCCAAGA AGGACGAGCC CAAGAGCGGC GAGGAGGCGC TCATCATCCC     240
113    CCCCGACGCC GTCGCGGTGG ACTGCAAGGA CCCAGATGAT GTGGTACCAG TTGGCCAAAG     300
114    AAGAGCCTGG TGTTGGTGCA TGTGCTTTGG ACTAGCATTT ATGCTTGCAG GTGTTATTCT     360
115    AGGAGGAGCA TACTTGTACA AATATTTTGC ACTTCAACCA GATGACGTGT ACTACTGTGG     420
116    AATAAAGTAC ATCAAAGATG ATGTCATCTT AAATGAGCCC TCTGCAGATG CCCCAGCTGC     480
117    TCTCTACCAG ACAATTGAAG AAAATATTAA AATCTTTGAA GAAGAAGAAG TTGAATTTAT     540
118    CAGTGTGCCT GTCCCAGAGT TTGCAGATAG TGATCCTGCC AACATTGTTC ATGACTTTAA     600
119    CAAGAAACTT ACAGCCTATT TAGATCTTAA CCTGGATAAG TGCTATGTGA TCCCTCTGAA     660
120    CACTTCCATT GTTATGCCAC CCAGAAACCT ACTGGAGTTA CTTATTAACA TCAAGGCTGG     720
121    AACCTATTTG CCTCAGTCCT ATCTGATTCA TGAGCACATG GTTATTACTG ATCGCATTGA     780
122    AAACATTGAT CACCTGGGTT TCTTTATTTA TCGACTGTGT CATGACAAGG AAACTTACAA     840
123    ACTGCAACGC AGAGAAACTA TTAAAGGTAT TCAGAAACGT GAAGCCAGCA ATTGTTTCGC     900
124    AATTCGGCAT TTTGAAAACA AATTTGCCGT GGAAACTTTA ATTTGTTCTT GAACAGTCAA     960
125    GAAAAACATT ATTGAGGAAA ATTAATATCA CAGCATAACC CCACCCTTTA CATTTTGTGC    1020
126    AGTGATTATT TTTTAAAGTC TTCTTTCATG TAAGTAGCAA ACAGGGCTTT ACTATCTTTT    1080
127    CATCTCATTA ATTCAATTAA AACCATTACC TTAA                               1114
128
129        (2) INFORMATION FOR SEQ ID NO:3:
130
131        (i) SEQUENCE CHARACTERISTICS:
132            (A) LENGTH: 263 amino acids
133            (B) TYPE: amino acid
134            (C) STRANDEDNESS: single
135            (D) TOPOLOGY: linear
136
137        (vii) IMMEDIATE SOURCE:
138            (A) LIBRARY: GenBank
139            (B) CLONE: 624778
140
141        (xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:
142
143    Met Val Lys Ile Ala Phe Asn Thr Pro Thr Ala Val Gln Lys Glu Glu
144     1               5                  10                  15
145    Ala Arg Gln Asp Ile Glu Ala Leu Val Ser Arg Thr Val Arg Ala Gln
146             20                  25                  30
147    Ile Leu Thr Gly Lys Glu Leu Arg Val Val Pro Gln Glu Lys Asp Gly
148             35                  40                  45
149    Ser Ser Gly Arg Cys Met Leu Thr Leu Leu Gly Leu Ser Phe Ile Leu
150         50                  55                  60
151    Ala Gly Leu Ile Val Gly Gly Ala Cys Ile Tyr Lys Tyr Phe Met Pro
152    65                  70                  75                  80
```

INPUT SET: S35114.raw

```
153    Lys Ser Thr Ile Tyr His Gly Glu Met Cys Phe Phe Asp Ser Glu Asp
154                    85                  90                  95
155    Pro Val Asn Ser Ile Pro Gly Gly Glu Pro Tyr Phe Leu Pro Val Thr
156                    100                 105                 110
157    Glu Glu Ala Asp Ile Arg Glu Asp Asp Asn Ile Ala Ile Ile Asp Val
158                    115                 120                 125
159    Pro Val Pro Ser Phe Ser Asp Ser Asp Pro Ala Ala Ile Ile His Asp
160                    130                 135                 140
161    Phe Glu Lys Gly Met Thr Ala Tyr Leu Asp Leu Leu Leu Gly Asn Cys
162    145                 150                 155                 160
163    Tyr Leu Met Pro Leu Asn Thr Ser Ile Val Met Thr Pro Lys Asn Leu
164                    165                 170                 175
165    Val Glu Leu Phe Gly Lys Leu Ala Ser Gly Lys Tyr Leu Pro His Thr
166                    180                 185                 190
167    Tyr Val Val Arg Glu Asp Leu Val Ala Val Glu Glu Ile Arg Asp Val
168                    195                 200                 205
169    Ser Asn Leu Gly Ile Phe Ile Tyr Gln Leu Cys Asn Asn Arg Lys Ser
170        210                 215                 220
171    Phe Arg Leu Arg Arg Arg Asp Leu Leu Leu Gly Phe Asn Lys Arg Ala
172    225                 230                 235                 240
173    Ile Asp Lys Cys Trp Lys Ile Arg His Phe Pro Asn Glu Phe Ile Val
174                    245                 250                 255
175    Glu Thr Lys Ile Cys Gln Glu
176                    260
177
178            (2) INFORMATION FOR SEQ ID NO:4:
179
180            (i) SEQUENCE CHARACTERISTICS:
181                (A) LENGTH: 1635 base pairs
182                (B) TYPE: nucleic acid
183                (C) STRANDEDNESS: single
184                (D) TOPOLOGY: linear
185
186            (vii) IMMEDIATE SOURCE:
187                (A) LIBRARY: GenBank
188                (B) CLONE: 624777
189
190            (xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:
191
192    GGGAGACCTG AGCTCGCTGC TGCCTGTGGA AGACTGGGAG AGGAGACACT AAGTGCTGCT      60
193    CAAGCAAGCG CGATCCTCTC CTCTTTCAAC CTGCAGCCCA AGATACTGAT TCGAGCCGCG     120
194    CCTTACCGCG CAGCCCGAAG ATTCACCATG GTGAAGATCG CCTTCAACAC CCCTACGGCG     180
195    GTGCAAAAGG AGGAGGCGCG GCAAGATATA GAGGCGCTCG TCAGTCGCAC TGTCCGAGCT     240
196    CAAATCCTGA CTGGCAAGGA GCTCAGAGTT GTCCCGCAGG AGAAAGATGG CTCATCTGGG     300
197    AGATGCATGC TTACTCTCCT AGGCCTCTCA TTCATCTTGG CAGGACTGAT TGTTGGTGGA     360
198    GCCTGCATTT ACAAGTACTT CATGCCCAAG AGCACCATTT ACCATGGTGA GATGTGCTTC     420
199    TTTGATTCTG AGGATCCTGT CAATTCCATT CCTGGAGGAG AGCCATACTT TCTGCCTGTG     480
200    ACTGAGGAGG CTGATATCCG TGAGGATGAC AACATTGCCA TCATTGATGT GCCTGTGCCC     540
201    AGTTTCTCTG ATAGCGATCC GGCGGCAATT ATTCACGACT TTGAGAAGGG AATGACTGCT     600
202    TACCTGGACT TGCTTTTGGG AAACTGTTAT CTGATGCCCC TCAATACTTC CATTGTTATG     660
203    ACTCCAAAGA ATCTGGTGGA ACTTTTTGGA AAACTGGCAA GTGGCAAGTA TTTGCCTCAT     720
204    ACTTATGTGG TTCGTGAAGA CCTGGTTGCT GTGGAAGAAA TTCGTGATGT TAGTAACCTT     780
205    GGTATTTTTA TTTACCAACT TTGCAACAAC CGAAAATCCT TCCGCCTTAG ACGCAGAGAC     840
```

RAW SEQUENCE LISTING
PATENT APPLICATION *US/09/207,161*

INPUT SET: S35114.raw

```
206   CTTCTGCTGG GTTTCAACAA GCGTGCCATT GACAAATGCT GGAAGATTAG ACACTTCCCC    900
207   AATGAATTTA TCGTTGAAAC CAAGATCTGT CAGGAGTGAA ATGTGACAGA TAAAGAGTAT    960
208   CCTTGATAAT AAGAAGTCAG GAACTTACCG TCTGACTTGG AAAATTGAAA TTGATGGGAT   1020
209   ACTCATGCTA TTTACTCATA CATTTACTCT ATTGCTTATA CTGGAAAAGG AAAGGGAAAG   1080
210   GGGGGAGAAA ACTACTAACC ACTGCAAGCG ATTGTCCAAT TCTACTTTAA TTGACATTGC   1140
211   TTGCTGTTTT CAACAAGTCA AATGATTATC TTTTCTCTTG AATTTATAGG GTTTAGATTT   1200
212   CTGAAAGCAG CATGAATGTG TCATCTTACC ATCCTGACAA TAAAGCCCAT CCTCTGGTTT   1260
213   TATTTAAAGC AAGCTCTTTC CAACATCACT TGGCTAGAGC ATGCTTTAAA TTTAAAATAT   1320
214   TTGAAATTTG TTTTTGACAT TTTTTTGTGT GAAACATGTC AAATCTCTTA CCATTCTTTG   1380
215   GTTTTCTTCT TTATTATGTT CAACTCTCCT GATTTCAGAA GTTACATTTT TGCATTTCTA   1440
216   TCAGGTGCTG TGTAACGAAT CTGACTGATA TGTGAACAAT CTTCATGAGG AAGCAATTTT   1500
217   TTACTCATGT AATGATTCTT TCTCACTGAT ATCTGTATTG TGAAATCCAC AGAACTGTAC   1560
218   AGGTGCTGAA TGCTGTAAGG AGTTCTGGTT GTATGAATTC TACAACCCTA TAATAAAGTT   1620
219   TACCGTATTC AATCA                                                   1635
220
221
```

SEQUENCE VERIFICATION REPORT
PATENT APPLICATION   *US/09/207,161*

INPUT SET: S35114.raw

Line   Error   Original Text